

Gibbs States and the Consistency of Local Density Matrices

Yi-Kai Liu

Computer Science and Engineering
University of California, San Diego
y9liu@cs.ucsd.edu

March 1, 2006

Abstract

Suppose we have an n -qubit system, and we are given a collection of local density matrices ρ_1, \dots, ρ_m , where each ρ_i describes some subset of the qubits. We say that ρ_1, \dots, ρ_m are “consistent” if there exists a global state σ (on all n qubits) whose reduced density matrices match ρ_1, \dots, ρ_m .

We prove the following result: if ρ_1, \dots, ρ_m are consistent with some state $\sigma \succ 0$, then they are also consistent with a state σ' of the form $\sigma' = (1/Z) \exp(M_1 + \dots + M_m)$, where each M_i is a Hermitian matrix acting on the same qubits as ρ_i , and Z is a normalizing factor. (This is known as a Gibbs state.) Actually, we show a more general result, on the consistency of a set of expectation values $\langle T_1 \rangle, \dots, \langle T_r \rangle$, where the observables T_1, \dots, T_r need not commute. This result was previously proved by Jaynes (1957) in the context of the maximum-entropy principle; here we provide a somewhat different proof, using properties of the partition function.

1 Introduction

Many-body systems have an intriguing property: under the right circumstances, local interactions can conspire to produce long-range or global effects. This behavior leads to phase transitions in statistical mechanics, and it also appears in combinatorial problems such as 3-SAT. If we consider quantum systems, the situation is more complicated, due to non-commuting measurements and the possibility of entanglement. This leads to new kinds of quantum phase transitions [1], and new examples such as the Local Hamiltonian problem [2].

A basic question in all of these examples is: if we know local information about various parts of a system, what can we say about the system as a whole? This paper gives one answer to this question, for quantum systems.

Suppose we have an n -qubit system, and we are given a collection of local density matrices ρ_1, \dots, ρ_m , where each ρ_i describes a subset $C_i \subseteq \{1, \dots, n\}$ of the qubits. We say that ρ_1, \dots, ρ_m are “consistent” if there exists a global state σ (on all n qubits) whose reduced density matrices match ρ_1, \dots, ρ_m ; in other words, for all $i = 1, \dots, m$, $\text{tr}_{\{1, \dots, n\} - C_i}(\sigma) = \rho_i$.

Clearly, if ρ_1, \dots, ρ_m are consistent, then whenever two density matrices ρ_i and ρ_j describe overlapping subsets of qubits ($C_i \cap C_j \neq \emptyset$), they must agree on the intersection $C_i \cap C_j$; that is, $\text{tr}_{C_i - (C_i \cap C_j)}(\rho_i) = \text{tr}_{C_j - (C_i \cap C_j)}(\rho_j)$. This gives a necessary condition for consistency.

However, the above condition is not sufficient to guarantee consistency. To see this, consider the following example: we have three qubits, and we are told that qubits 1 and 2 are in the Bell state $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, and qubits 2 and 3 are also in the same state $|\Phi^+\rangle$. More formally, let $\rho_A = |\Phi^+\rangle\langle\Phi^+|$, $A = \{1, 2\}$, and let $\rho_B = |\Phi^+\rangle\langle\Phi^+|$, $B = \{2, 3\}$. In this case, ρ_A and ρ_B both agree on qubit 2, since $\text{tr}_1(\rho_A) = I/2 = \text{tr}_3(\rho_B)$. But there is no state σ on all three qubits such that $\text{tr}_3(\sigma) = \rho_A$ and $\text{tr}_1(\sigma) = \rho_B$; one way to see this is to apply the strong subadditivity inequality, $S(1, 2, 3) + S(2) \leq S(1, 2) + S(2, 3)$.

Thus the consistency of ρ_1, \dots, ρ_m would seem to be a more subtle question. We prove the following result:

Theorem 1 *If ρ_1, \dots, ρ_m are consistent with some state $\sigma \succ 0$, then they are also consistent with a state σ' of the form $\sigma' = (1/Z) \exp(M_1 + \dots + M_m)$, where each M_i is a Hermitian matrix acting on the qubits in C_i , and $Z = \text{tr}(\exp(M_1 + \dots + M_m))$.*

Here, $\sigma \succ 0$ means that σ is a positive definite matrix. The state σ' is known as a Gibbs state.

Essentially, this result says that a Gibbs state σ' can simulate an arbitrary state $\sigma \succ 0$, with respect to an observer who can only access subsets C_1, \dots, C_m of the qubits. For example, consider a physical system with lo-

cal interactions, described by a Hamiltonian H . It is easy to see that the ground state of H can be approximated by $\eta = (1/Z) \exp(-\beta H)$, for β large; and since H is a sum of local terms, η is a Gibbs state. Our result extends this simple observation to a much more general setting.

Actually, we prove the following more general result: Consider a finite quantum system, and let T_1, \dots, T_r be observables (Hermitian matrices). Without loss of generality, assume that the collection of matrices I, T_1, \dots, T_r is linearly independent (over \mathbb{R}). We say that a state ρ has expectation values t_1, \dots, t_r if $\text{tr}(T_i \rho) = t_i$ for all $i = 1, \dots, r$.

Theorem 2 *If there exists some state $\rho \succ 0$ which has expectation values t_1, \dots, t_r , then there exists a state ρ' which has the same expectation values t_1, \dots, t_r , and is of the form $\rho' = (1/Z) \exp(\theta_1 T_1 + \dots + \theta_r T_r)$, where $\theta_1, \dots, \theta_r \in \mathbb{R}$.*

This statement holds even when the observables T_1, \dots, T_r do not commute.

This result was previously proved by Jaynes, as part of the maximum entropy principle in statistical mechanics [3, 4]. Jaynes showed that the Gibbs state ρ' is the state which maximizes the entropy $S(\rho) = -\text{tr}(\rho \log \rho)$ subject to the constraints $\langle T_i \rangle = t_i$; implicitly, he also showed that the Gibbs state ρ' is always feasible, in the sense that it can produce the same expectation values $\langle T_i \rangle$ as an arbitrary state $\rho \succ 0$.

However, Jaynes' motivation was somewhat different from ours. Jaynes was interested in statistical mechanics, which deals with large systems with many degrees of freedom and only a few constraints. Feasibility is not usually a concern in such cases, while the maximum-entropy property is crucial in making plausible inferences about the "true" state of the system.

In this paper, we focus on finite quantum systems, with many non-commuting constraints; we are interested in the relationship between local constraints and the global state of the system. For us, feasibility of the Gibbs state is important, since it is possible for the system to become overdetermined. Statistical inference is less important, because the systems we study are small enough that their state can be completely determined (at least in principle). Rather than viewing this as an inference problem, we can speak directly about what states are allowed under a given set of constraints.

Finally, we prove our result using a technique which is different from Jaynes. Jaynes used the Lagrange dual of the entropy-maximization problem, while we use some analytic properties of the partition function. Our analysis bears some resemblance to classical results on exponential families in statistics [5]—although the technical

details are quite different. Our proof also contains some geometric intuition which may be of interest.

2 Proofs of our results

First, we will review some useful facts about the partition function for a Gibbs state. Then we will prove theorem 2, and obtain theorem 1 as a special case.

2.1 The partition function

Recall the situation described in theorem 2: we have a finite quantum system, and observables T_1, \dots, T_r , such that I, T_1, \dots, T_r are linearly independent (over \mathbb{R}). We are interested in states of the form

$$\rho(\theta) = \exp(\theta_1 T_1 + \dots + \theta_r T_r) / Z(\theta), \quad \theta \in \mathbb{R}^r,$$

where $Z(\theta) = \text{tr}(\exp(\theta_1 T_1 + \dots + \theta_r T_r))$. $Z(\theta)$ is called the partition function, and we also define the log partition function $\psi(\theta) = \log Z(\theta)$.

Note that, in the above definition, we can translate each observable T_i by a multiple of the identity, without changing the state $\rho(\theta)$. More precisely, if we define new observables $P_i = T_i + \lambda_i I$, with $\lambda_i \in \mathbb{R}$, we have that:

$$\frac{\exp(\theta_1 P_1 + \dots + \theta_r P_r)}{\text{tr}(\exp(\theta_1 P_1 + \dots + \theta_r P_r))} = \frac{\exp(\theta_1 T_1 + \dots + \theta_r T_r)}{\text{tr}(\exp(\theta_1 T_1 + \dots + \theta_r T_r))}.$$

Using subscripts T and P to denote the two sets of observables, we arrive at the same state, $\rho_P(\theta) = \rho_T(\theta)$, although the partition functions are different, $Z_P(\theta) \neq Z_T(\theta)$.

The log partition function ψ has some nice analytic properties: it is convex, and its derivatives encode the expectation values of the observables T_i . We briefly sketch these results, which can be found in quantum statistical mechanics [4], as well as quantum information geometry [6].

Proposition 3 *ψ is convex on \mathbb{R}^r .*

Proof sketch: This follows from some facts in matrix analysis [7]. First, the Golden-Thompson inequality: If A and B are Hermitian matrices, then

$$\text{tr}(\exp(A + B)) \leq \text{tr}(\exp(A) \exp(B)).$$

Next, a matrix version of Hölder's inequality: For any matrix A , define the Frobenius or Hilbert-Schmidt norm to be $\|A\|_2 = (\text{tr}(A^\dagger A))^{1/2}$. Also, let $|A|$ denote the unique positive semidefinite square root of $A^\dagger A$. Then we have that, for all square matrices A and B ,

$$\|AB\|_2 \leq \| |A|^p \|_2^{1/p} \| |B|^q \|_2^{1/q},$$

for $\frac{1}{p} + \frac{1}{q} = 1$, $p > 1$. \square

Proposition 4 ψ is differentiable on \mathbb{R}^r , and

$$\frac{\partial \psi}{\partial \theta_i} = \text{tr}(T_i \rho(\theta)) = \langle T_i \rangle.$$

Proof sketch: Use “parameter differentiation” [8]: If H is a Hermitian matrix which depends on a parameter λ , and $\partial H / \partial \lambda$ and $\partial^2 H / \partial \lambda^2$ exist and are continuous, then $\partial(\exp(H)) / \partial \lambda$ exists and is equal to

$$\frac{\partial}{\partial \lambda} \exp(H) = \int_0^1 \exp((1-u)H) \frac{\partial H}{\partial \lambda} \exp(uH) du. \quad \square$$

2.2 Proof of theorem 2

Proof: We are given expectation values t_1, \dots, t_r , and we want to find a state

$$\rho'(\theta) = \exp(\theta_1 T_1 + \dots + \theta_r T_r) / Z'(\theta)$$

that has these expectation values. (Here, $Z'(\theta)$ is the partition function, and $\psi'(\theta) = \log Z'(\theta)$ is the log partition function.) By translating the observables T_i , we can assume that $t_i = 0$, for all $i = 1, \dots, r$. We can now restate the problem in terms of the log partition function: we are looking for some $\theta \in \mathbb{R}^r$ such that $\nabla \psi'(\theta) = 0$.

We know there exists a state $\rho \succ 0$ which has the desired expectation values t_1, \dots, t_r . Now choose some observables U_1, \dots, U_s , such that the set $\{I, T_1, \dots, T_r, U_1, \dots, U_s\}$ is complete and linearly independent (in other words, any 2^n -dimensional Hermitian matrix can be written uniquely as a real linear combination of the matrices in this set). Let u_1, \dots, u_s be the expectation values of ρ for the observables U_1, \dots, U_s ; that is, $u_i = \text{tr}(U_i \rho)$. By translating the U_i , we can assume that $u_i = 0$, for all $i = 1, \dots, s$.

We will consider states of the form

$$\rho(\theta, \phi) = \exp(\theta_1 T_1 + \dots + \theta_r T_r + \phi_1 U_1 + \dots + \phi_s U_s) / Z(\theta, \phi).$$

(Here, $Z(\theta, \phi)$ is the partition function, and $\psi(\theta, \phi) = \log Z(\theta, \phi)$ is the log partition function.) Completeness of the T_i and the U_i implies that we can write ρ in the form $\rho = \rho(\theta, \phi)$ for some $(\theta, \phi) \in \mathbb{R}^{r+s}$. This implies that $\nabla \psi(\theta, \phi) = 0$ for some $(\theta, \phi) \in \mathbb{R}^{r+s}$.

Furthermore, we claim that there is a unique point (θ, ϕ) such that $\rho(\theta, \phi)$ has the expectation values t_i and u_i . This is because the expectation values t_i and u_i uniquely determine the state ρ , and setting $\rho = \rho(\theta, \phi)$ uniquely determines the values of θ and ϕ . This in turn follows from the completeness and linear independence of

the T_i and the U_i . So we conclude that $\nabla \psi(\theta, \phi) = 0$ at exactly one point (θ, ϕ) .

To complete the proof, we will carry out the following plan: we will show that $\psi(\theta, \phi) \rightarrow \infty$ as $\|\theta, \phi\| \rightarrow \infty$, where $\|\theta, \phi\|$ denotes the norm of the vector (θ, ϕ) . This implies that $\psi'(\theta) \rightarrow \infty$ as $\|\theta\| \rightarrow \infty$; and hence $\nabla \psi'(\theta) = 0$ for some $\theta \in \mathbb{R}^r$. (See figure 1 for a simple example that shows the geometric intuition for the proof.)

Let (θ_0, ϕ_0) be the unique point where $\nabla \psi$ vanishes. We claim that (θ_0, ϕ_0) is the unique global minimum of ψ . [Since ψ is convex (proposition 3), it follows that ψ is bounded below, and (θ_0, ϕ_0) is a global minimum. Also, ψ is differentiable everywhere on the domain \mathbb{R}^{r+s} , which has no boundaries (proposition 4); so any extremum (θ, ϕ) must satisfy $\nabla \psi(\theta, \phi) = 0$. But this happens only at (θ_0, ϕ_0) , and so (θ_0, ϕ_0) is the unique global minimum.]

Let S be the set of all unit vectors in \mathbb{R}^{r+s} . Define the function $f(\nu, z) = \psi((\theta_0, \phi_0) + z\nu)$, for $\nu \in S$, and $z \in \mathbb{R}$. Say we fix $z = 1$. We claim that there exists some $b > 0$ such that, for all ν , $f(\nu, 1) \geq \psi(\theta_0, \phi_0) + b$. [Since (θ_0, ϕ_0) is the unique global minimum, we have that $f(\nu, 1) > \psi(\theta_0, \phi_0)$, for all ν . Moreover, $f(\nu, 1)$ is a continuous function of ν , and S is a compact set, hence its image $f(S, 1)$ is compact. Hence $f(\nu, 1)$ must be bounded away from $\psi(\theta_0, \phi_0)$, for all ν .]

Next we claim that, for all ν , and for all $z \geq 1$, $(\partial f / \partial z)(\nu, z) \geq b$. [Fix any ν . $f(\nu, z)$ is a differentiable function of z , so by the mean value theorem, there exists some $z \in (0, 1)$ such that $(\partial f / \partial z)(\nu, z) = f(\nu, 1) - f(\nu, 0) \geq b$. In addition, since ψ is convex, $(\partial f / \partial z)(\nu, z)$ is nondecreasing in z . This proves the claim.]

Now, say we are given some (θ, ϕ) , and assume that $\|(\theta, \phi) - (\theta_0, \phi_0)\| \geq 1$. We can write (θ, ϕ) in the form

$$(\theta, \phi) = (\theta_0, \phi_0) + \|(\theta, \phi) - (\theta_0, \phi_0)\| \nu,$$

for some unit vector $\nu \in S$. Then we have:

$$\begin{aligned} \psi(\theta, \phi) &= f(\nu, \|(\theta, \phi) - (\theta_0, \phi_0)\|) \\ &= f(\nu, 1) + \int_1^{\|(\theta, \phi) - (\theta_0, \phi_0)\|} (\partial f / \partial z)(\nu, z) dz \\ &\geq \psi(\theta_0, \phi_0) + b + b(\|(\theta, \phi) - (\theta_0, \phi_0)\| - 1) \\ &= \psi(\theta_0, \phi_0) + b\|(\theta, \phi) - (\theta_0, \phi_0)\|. \end{aligned}$$

From this, we conclude that $\psi(\theta, \phi) \rightarrow \infty$ as $\|\theta, \phi\| \rightarrow \infty$.

Notice that the partition functions for $\rho'(\theta)$ and $\rho(\theta, \phi)$ are related:

$$\psi'(\theta) = \psi(\theta, 0).$$

Hence, $\psi'(\theta) \rightarrow \infty$ as $\|\theta\| \rightarrow \infty$.

We will use the following fact: if $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous, and $f(x) \rightarrow \infty$ as $\|x\| \rightarrow \infty$, then f is bounded

below, and f attains its minimum at some point $x_* \in \mathbb{R}^n$. [To see this, let $S = \{x \in \mathbb{R}^n \mid f(x) \leq \alpha\}$, choosing α large enough that $S \neq \emptyset$. Note that S is bounded; otherwise, there would exist a sequence $\{x_i\}$ such that $\|x_i\| \rightarrow \infty$ and $f(x_i) \leq \alpha$, a contradiction. Also, note that S is closed; this is because the interval $(-\infty, \alpha]$ is closed, and f is continuous. So we have that S is compact. This implies that $f(S)$ is compact. Hence $f(S)$ is closed and bounded; also note that $f(S) \neq \emptyset$. This implies that f is bounded below, and attains its minimum.]

From this, we conclude that ψ' attains its minimum at some point $\theta_* \in \mathbb{R}^r$. \mathbb{R}^r has no boundaries, and ψ' is differentiable everywhere on \mathbb{R}^r , so it follows that $\nabla \psi'(\theta_*) = 0$. This completes the proof. \square

2.3 Proof of theorem 1

Proof: We will obtain theorem 1 as a special case of theorem 2. The basic idea is that specifying the local density matrices ρ_1, \dots, ρ_m is equivalent to specifying the expectation values of all Pauli matrices on the subsets C_1, \dots, C_m .

Let X, Y and Z denote the Pauli matrices for a single qubit, and define $\mathcal{P} = \{I, X, Y, Z\}$. We can construct n -qubit Pauli matrices by taking tensor products $P = P_1 \otimes \dots \otimes P_n \in \mathcal{P}^{\otimes n}$. Any 2^n -dimensional Hermitian matrix can be written as a real linear combination of n -qubit Pauli matrices. Furthermore, the n -qubit Pauli matrices are orthogonal with respect to the Hilbert-Schmidt inner product: $\text{tr}(P^\dagger Q) = 2^n$ if $P = Q$, and 0 otherwise.

We make the following claim: Let σ be a density matrix on n qubits, and let ρ be a density matrix on a subset of the qubits $C \subseteq \{1, \dots, n\}$, with $|C| = k$. We claim that $\text{tr}_{\{1, \dots, n\}-C}(\sigma) = \rho$, if and only if, for all Pauli matrices P on the subset C , $\text{tr}((P \otimes I)\sigma) = \text{tr}(P\rho)$. (Notation: we write n -qubit Pauli matrices in the form $P \otimes Q$, where P acts on the subset C , and Q acts on the rest of the qubits.)

The (\Rightarrow) direction is obvious, but we need to show (\Leftarrow) . We write σ and ρ as linear combinations of Pauli matrices, with real coefficients $\beta_{(P \otimes Q)}$ and α_P :

$$\begin{aligned}\sigma &= \sum_{(P \otimes Q) \in \mathcal{P}^{\otimes n}} \beta_{(P \otimes Q)} P \otimes Q \\ \rho &= \sum_{P \in \mathcal{P}^{\otimes k}} \alpha_P P.\end{aligned}$$

We know that, for all Pauli matrices P on the subset C , $\text{tr}((P \otimes I)\sigma) = 2^n \beta_{(P \otimes I)} = \text{tr}(P\rho) = 2^k \alpha_P$. But this

implies:

$$\begin{aligned}\text{tr}_{\{1, \dots, n\}-C}(\sigma) &= \sum_{P \in \mathcal{P}^{\otimes k}} 2^{n-k} \beta_{(P \otimes I)} P \\ &= \sum_{P \in \mathcal{P}^{\otimes k}} \alpha_P P = \rho,\end{aligned}$$

which proves the claim.

Thus, theorem 1 is a special case of theorem 2, where the observables T_1, \dots, T_r consist of all the Pauli matrices on the subsets C_1, \dots, C_m . \square

Acknowledgements: I am grateful to Dorit Aharonov, Chris Fuchs and David Meyer for helpful discussions about this work. Funded by an ARO/NSA Quantum Computing Graduate Research Fellowship.

References

- [1] S. Sachdev, *Quantum Phase Transitions*, Cambridge University Press, 2000.
- [2] D. Aharonov and T. Naveh, “Quantum NP - A Survey,” Arxiv: quant-ph/0210077.
- [3] E.T. Jaynes, “Information Theory and Statistical Mechanics II,” Phys. Rev. (2) 108 (1957), pp.171-190.
- [4] E.T. Jaynes, “Information Theory and Statistical Mechanics,” lectures at Brandeis in 1962. Reprinted in *E.T. Jaynes: Papers on Probability, Statistics and Statistical Physics*, R.D. Rosenkrantz (ed.), D. Reidel Publishing Company, 1983.
- [5] L.D. Brown, *Fundamentals of Statistical Exponential Families*, Institute of Mathematical Statistics, 1986.
- [6] R.S. Ingarden, H. Janyszek, A. Kossakowski and T. Kawaguchi, “Information geometry of quantum statistical systems,” Tensor (N.S.) 37 (1982), pp.105-111.
- [7] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics, Springer-Verlag, 1997.
- [8] R.M. Wilcox, “Exponential Operators and Parameter Differentiation in Quantum Physics,” J. Mathematical Physics 8 (1967), pp.962-982.

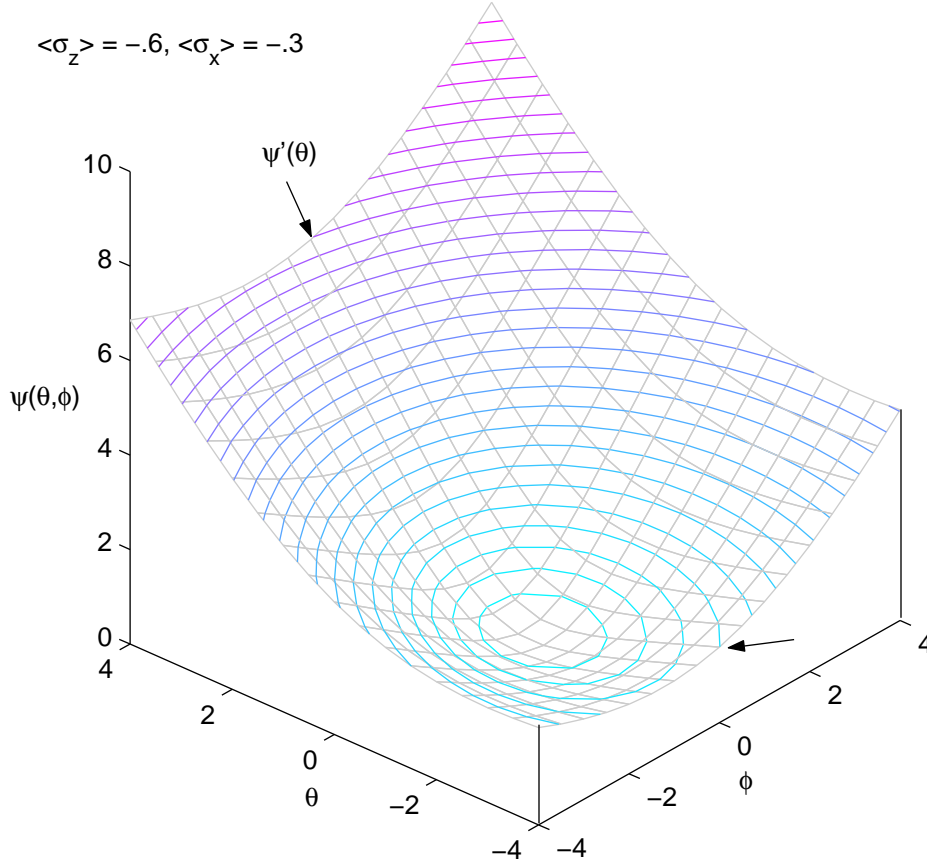


Figure 1: A single-qubit example. We want to find a Gibbs state ρ' that satisfies $\langle \sigma_z \rangle = -0.6$; we have one observable $T = \sigma_z + 0.6$. We know there exists some state $\rho \succ 0$ that satisfies $\langle \sigma_z \rangle = -0.6$; in this case, ρ also satisfies $\langle \sigma_x \rangle = -0.3$, and we let $U = \sigma_x + 0.3$ play the role of the “extra” observables. As the graph shows, $\nabla \psi(\theta, \phi)$ vanishes at exactly one point; $\psi'(\theta) = \psi(\theta, 0)$; and $\nabla \psi'(\theta)$ vanishes for some θ .